



**HAL**  
open science

## **Modélisation incertaine à partir de mesures non-reproductibles. Application à la comparaison de pression exercée par des matelas**

Cyprien Neverov, Chihab Khnifass, Papa Beye, Nicolas Sutton-Charani, Abdelhak Imoussaten, W Fagart, M. Blot, A. Dupeyron

### ► To cite this version:

Cyprien Neverov, Chihab Khnifass, Papa Beye, Nicolas Sutton-Charani, Abdelhak Imoussaten, et al.. Modélisation incertaine à partir de mesures non-reproductibles. Application à la comparaison de pression exercée par des matelas. LFA 2019 - 28ème Rencontres Francophones sur la Logique Floue et ses Applications, Nov 2019, Alès, France. hal-02294372

**HAL Id: hal-02294372**

**<https://hal.science/hal-02294372>**

Submitted on 23 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Modélisation incertaine à partir de mesures non-reproductibles. Application à la comparaison de pression exercée par des matelas

## Uncertain modelling from repeated measurements of non-reproducible experiments. Application to the mattresses pressure comparison

C. Neverov<sup>1</sup>

C. Khnifass<sup>1</sup>  
W. Fagart<sup>2</sup>

P. Beye<sup>1</sup>  
M. Blot<sup>2</sup>

N. Sutton-Charani<sup>1</sup>  
A. Dupeyron<sup>2</sup>

A. Imoussaten<sup>1</sup>

LGI2P, IMT Mines-Alès<sup>1</sup>

CHU de Nîmes<sup>2</sup>

6 avenue de Clavières F-30 319 Alès Cedex<sup>1</sup>,

4 Rue du Professeur Robert Debré, 30029 Nîmes<sup>2</sup>,

{cyprien.neverov, chihab.khnifass, papa.beye}@mines-ales.org,  
{nicolas.sutton-charani, abdelhak.imoussaten}@mines-ales.fr

{willy.fagart, mylene.blot}@chu-nimes.fr,  
arnaud.dupeyron@umontpellier.fr

### Résumé :

Cet article concerne la modélisation de données incertaines à partir de mesures répétées mais non-reproductibles. Différentes approches issues des théories modernes de l'incertain sont considérées. Plus précisément, l'article s'intéresse à diverses extensions de tests statistiques non-paramétriques aux données incertaines. Dans ce travail préliminaire, les modèles proposés sont mis en œuvre pour étendre une étude de comparaison de pression exercée par deux types de matelas sur les patients du CHU de Nîmes.

### Mots-clés :

Tests statistiques, données incertaines, escarres, p-valeur, non-reproductibilité.

### Abstract:

This paper presents the modeling of uncertain data from repeated but non-reproducible measurements. Several approaches related to modern uncertainty theories are considered. More precisely, different extensions of non-parametric statistical tests to uncertain data are investigated. In this preliminary work, the proposed models are implemented in order to extend a study of pressure comparison between two types of mattress realized at the university hospital of Nîmes.

### Keywords:

Statistical tests, uncertain data, pressure ulcer, p-value, non-reproducible.

## 1 Introduction

Une escarre est une lésion cutanée d'origine ischémique liée à une compression des tissus mous entre un plan dur et les saillies osseuses. Dans les services impliquant des hospitalisations de longue durée et chez les personnes âgées, la pression exercée par les matelas sur le sacrum peut être suffisante pour

provoquer des escarres. En milieu hospitalier en général, et plus particulièrement dans les services de rééducation, de réadaptation et de gériatrie, il s'agit d'un problème majeur. Les survenues d'escarres entraînent des périodes d'hospitalisation rallongées, ce qui augmente considérablement les prises en charge et représente donc un enjeu économique de taille. Les personnes atteintes de maladies chroniques qui imposent une position allongée sur un lit ou assise dans un fauteuil roulant sont les plus touchées. En 2015, le Centre d'Évaluation des Dispositifs Médicaux Handicap (CEDM-H) du Centre Hospitalier Universitaire (CHU) de Nîmes a mené une étude qui avait pour but de comparer deux types de matelas en termes de provocation d'escarres chez les patients hospitalisés (et donc alités) sur de longues durées.

Pour comparer les deux types de matelas, les chercheurs-statisticiens du BESPIM (Laboratoire de Biostatistique, Épidémiologie clinique, Santé Publique, Innovation et Méthodologie) rattaché au CHU ont constitué une cohorte de 60 patients suivie sur environ un mois. Le critère principal était la pression exercée par les matelas sur le patient au niveau de la zone sacrée. De manière à évaluer la fiabilité des données, les mesures réalisées selon différents protocoles expérimentaux ont été répétées cha-

cune deux fois pour mesurer la reproductibilité des mesures. Les données obtenues présentent une variabilité significative entre répétitions. Cette non-reproductibilité, et donc non-fiabilité des données a empêché les statisticiens en charge de l'étude de conclure quant à la comparaison entre les matelas.

Cet article propose une approche issue des théories de l'incertain permettant le traitement de telles données imparfaites. De nombreuses approches *soft* concernent l'extension de méthodes classiques d'apprentissage automatique [4, 16]. D'autres travaux moins nombreuses, sont dédiés à l'extension de tests statistiques usuels aux données incertaines [5, 18]. Dans des contextes de données peu nombreuses, de mauvaise qualité ou pour lesquelles il est difficile de faire des hypothèses (de normalité, d'égalité de variance, etc), les tests statistiques *non-paramétriques* sont généralement employés de manière à conclure quant à la significativité de différences observées au sein d'échantillons de données. Les différents tests non-paramétriques étudiés dans cet article sont connus sous les nom de test de Kolmogorov-Smirnov, test de Mann-Whitney, test des signes et test de Wilcoxon. Ils permettent tous la comparaison d'échantillons (en termes de distribution ou de médiane).

L'approche explorée dans ce papier consiste, à partir des données précises du CHU issues de mesures répétées et présentant une variabilité importante, à construire des modèles d'incertitude traduisant finement cette non-reproductibilité afin d'appliquer certaines extensions de tests statistiques aux données incertaines ou imprécises (*tests incertains*) de manière à conclure quant à la comparaison de matelas. Le but de notre travail est donc triple : consolider la liste des tests *incertains* non-paramétriques disponibles dans la littérature, proposer des approches de modélisation incertaines pour les données (précises) répétées mais non-reproductibles et enfin aider les statisticiens du CHU à conclure quant à la comparaison des matelas pour le CEDM-H.

Après avoir détaillé la problématique *matelas* du CHU en Section 2, un rappel concernant les notions de base des théories de l'incertain est effectué en Section 3, un bref état de l'art sur les tests *incertains* (applicables aux données incertaines, imprécises) non-paramétriques est proposé en Section 4, les différents modèles d'incertitude applicables aux données issues de mesures répétées sont présentés en Section 5 et implémentés de façon à mettre en œuvre les tests incertains adéquats pour répondre à la problématique *matelas* du CHU en Section 6.

## 2 Problématique - CHU

Les mécanismes physiologiques provoquant l'apparition d'escarres sont encore assez mal connus par les experts, il est cependant généralement admis par les praticiens que c'est principalement au niveau du sacrum des patients qu'elles surviennent. Lors de la planification de l'expérience, le personnel du CEDM-H a défini le critère principal de comparaison des deux types de matelas par les pics de pression exercée sur la zone sacrée des patients. La cohorte de patients était répartie sur 3 centres (CHU Nîmes - EHPAD La Chimotaie - EHPAD Notre Dame des Pins), chaque centre effectuant les mesures sur 20 patients pour chacun des deux types de matelas. Les pics et surfaces de pression sont mesurés grâce à une nappe de capteurs disposée sur les matelas.

Pour un patient donné, les mesures sont effectuées sur une période de deux jours, un matelas par jour. Et sur chaque matelas, deux mesures avec dossier en position  $0^\circ$  (position *sommeil*) et deux mesures avec  $30^\circ$  de surélévation du dossier (position *lecture-TV*). Chaque donnée (précise) correspond à une moyenne calculée sur 30 minutes du pic de pression spatial au niveau de la zone sacrée.

Chaque mesure du tableau étant réalisée sur un laps de temps de trente minutes, les patients doivent rester immobiles sur le matelas pendant la totalité de cette durée. Cette immobilité représente une contrainte non-négligeable, sur-

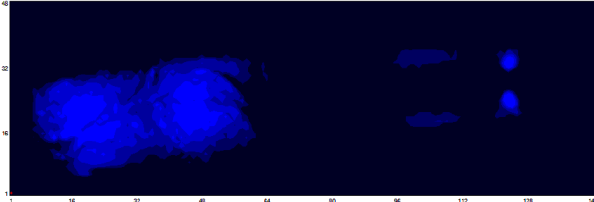


Figure 1 – Exemple de carte de pression

tout pour des populations souvent assez âgées. Quinze minutes de repos sont donc prévues entre chaque mesure pour permettre aux patients de se reposer et au matelas de revenir dans leur conditions initiales.

Pendant les 30 minutes de mesure, la nappe enregistre les valeurs de pression de chaque capteur et génère une carte de chaleur ou *heat-map* (comme illustré sur la Figure 1) sur laquelle on peut voir la silhouette du patient allongé. Pour réaliser la mesure du critère principal, le logiciel mesure chaque seconde le pic spatial de pression dans un rectangle qui englobe la zone sacrée du patient. La valeur finale est la moyenne de ces mesures calculée sur une demi-heure.

La mesure de reproductibilité, ou indice de concordance utilisée dans l'étude statistique réalisée par le CHU est la Corrélation Intra-Classe  $\rho$  définie dans un contexte d'analyse de variance (ANOVA) :

$$\rho = \frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_{\epsilon}^2} \quad (1)$$

où  $\sigma_{\alpha}^2$  et  $\sigma_{\epsilon}^2$  sont respectivement la variance de l'effet *groupe* et celle de l'effet *bruit*.

Des données sont généralement considérées répétables lorsqu'on a  $\rho \geq 0.95$ . Dans le cas de l'étude du CHU, les valeurs de  $\rho$  varient entre 0.5 et 0.73 ce qui est trop faible pour appliquer les outils de statistique conventionnels.

Pour un même triplet (patient, matelas, position) les deux séries temporelles (données brutes de pression avant moyennage) calculées sur trente minutes peuvent être totalement différentes et présentent souvent des décalages de niveaux. Cela s'explique probablement par

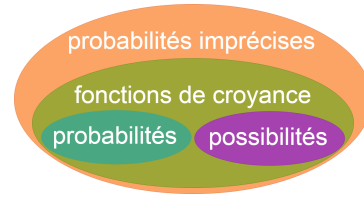


Figure 2 – Théories de l'incertain

le fait que les patients ne se placent souvent pas exactement de la même manière sur un matelas, les capteurs sollicités peuvent ainsi différer et il n'est donc pas invraisemblable d'observer de la variabilité au sein de données censées mesurer la même chose. La section suivante présente certains outils pouvant rendre compte de telles imperfections.

### 3 Rappels sur les théories de l'incertain

Au sein du *Soft Computing*, les théories modernes de l'incertain permettent d'intégrer l'imprécision et/ou l'incertitude inhérente aux données issues d'applications de la plupart des domaines tels que la biologie, la médecine, les sciences humaines, etc. La robustesse, la complexité et la faisabilité sont des critères importants pendant l'implémentation des modèles *incertains* ou *soft*. Pour atteindre notre objectif final, qui est d'aider le CHU à faire un choix de matelas, nous devons d'abord transformer les données *matelas* de manière à ce qu'elles rendent compte de l'imprécision et l'incertitude générées par les capteurs.

Quelques notions de base nécessaires aux modèles proposés en Section 5 sont ici introduites. Le terme "théories de l'incertain" désigne l'ensemble formé par les théories des probabilités [1, 13], des possibilités [21, 7], des fonctions de croyance [3, 15] et des probabilités imprécises [19]. De nombreux liens ont été explicités entre ces théories, une vision (un peu simplifiée) peut être résumée par la Figure 2.

Les modèles de données incertaines proposés dans ce travail seront de trois formes : des in-

tervalles, des distributions de possibilités et des fonctions de croyance.

**Possibilités.** Les sous-ensembles flous sont caractérisés par un degré d'appartenance qui, au lieu d'être binaire comme en théorie des ensembles, est à valeurs dans  $[0, 1]$ . La théorie des possibilités peut être considérée comme le pan incertain de la théorie des sous-ensembles flous. On considère deux mesures d'incertitude duales : la nécessité et la possibilité. Ces dernières sont définies relativement à un sous-ensemble flou.

Nous utilisons un cas particulier de mesure de possibilité : la *distribution trapezoidale* ou *intervalle* flou. Elle est caractérisée par le quadruplet de nombres réels  $(a, b, c, d)$ ,  $a < b < c < d$  tels que l'appartenance croît de 0 à 1 entre  $a$  et  $b$ , reste constante et égale à 1 entre  $b$  et  $c$  puis décroît de 1 à 0 entre  $c$  et  $d$ . La figure 3 illustre ce type de distribution de possibilité.

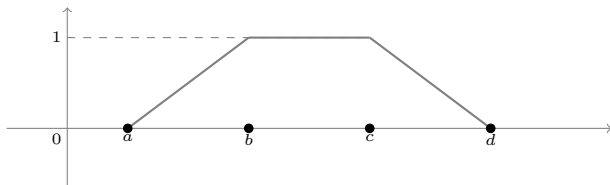


Figure 3 – Distribution trapezoidale

**Fonctions de croyance.** Les possibilités sont un cas particulier de fonction de croyance qui couvrent donc un spectre de modèles plus large, ce qui implique aussi une plus grande complexité que celle des possibilités. Ici encore, deux mesures duales d'incertitude sont considérées : la croyance ( $Bel$ ) et la plausibilité ( $Pl$ ). La désignation anglophone de cette théorie, *evidence theory*, traduit bien la philosophie des fonctions de croyance : se baser uniquement sur l'information dont on dispose.

Une *masse de croyance* est définie de manière à représenter le niveau de connaissance relatif à toutes les sous-parties de l'espace  $\Omega$  considéré, elle est donc définie de  $2^\Omega$  telle que  $\sum_{A \in 2^\Omega} m(A) = 1$  et on pose généralement  $m(\emptyset) = 0$ . Les éléments de  $2^\Omega$  ayant une masse

Tableau 1 – Tests incertains non-paramétriques

Type de test	Données	Référence
Kolmogorov-Smirnov	intervalles	Destercke-Strauss [6]
Mann-Whitney	floues	Denœux et al. [5]
Mann-Whitney	floues	Hesamian-Taheri [11]
Mann-Whitney	intervalles	Perolat et al.[2]
test des signes	floues	Hesamian-Chachi [10]
Wilcoxon	floues	Vaidyanathan [17]
Wilcoxon	floues	Taheri-Hesamian [9]
Wilcoxon	floues	Kahraman et al. [12]

strictement positive sont appelés les éléments focaux de  $m$ . Les fonctions de croyance et de plausibilité sont ainsi définies : pour  $A \subseteq \Omega$ ,  $Bel(A) = \sum_{B \subseteq A} m(B)$  et  $Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$ .

Les possibilités sont des fonctions de croyance pour lesquelles tous les éléments focaux sont emboîtés.

## 4 Tests statistiques non-paramétriques incertains ou imprécis

Les théories modernes de l'incertain permettent une prise en compte fine des incertitudes et imprécisions inhérentes aux données. Lorsqu'en plus de ces imperfections les données sont peu nombreuses, émettre des hypothèses quant à leur distribution se révèle illusoire.

Dans ce contexte de données peu nombreuses, la mise en œuvre de tests dits *non-paramétriques* est d'usage. Bien qu'un peu moins puissants que les tests paramétriques, presque aucune hypothèse n'est faite quant au modèle génératif (inconnu) des données. Denœux et al. justifient l'utilisation des tests non-paramétriques dans [5] par le fait qu'il est difficile de parler de normalité ou d'égalité de variance pour des données de mauvaise qualité.

Pour comparer les distributions de deux échantillons à l'aide d'un test non-paramétrique, les tests de Kolmogorov-Smirnov, de Mann-Whitney, des signes et de Wilcoxon sont des solutions classiques.

Différentes extensions de tests non-paramétriques aux données incertaines existent dans la littérature, le tableau 1 en rassemble les

principales. Ces méthodes peuvent prendre en entrée deux types de données correspondant à deux niveaux de représentation de l'incertitude différents : l'intervalle ou le sous-ensemble flou, le dernier étant plus riche en terme d'information mais plus complexe d'un point de vue calculatoire.

Ces tests incertains donnent généralement lieu à des p-valeurs imprécises (sauf pour [17]). La complexité des p-valeurs dépend de la complexité des données : un test sur données intervalles donnera généralement une p-valeur sous forme d'intervalle alors qu'une p-valeur floue sera obtenue à partir d'un test sur données floues.

Un point commun à presque tous ces travaux est qu'ils aboutissent potentiellement à trois conclusions : le rejet de l'hypothèse nulle  $H_0$ , le non-rejet de  $H_0$ , ou l'impossibilité de conclure de manière significative (en absence de données supplémentaires). Rappelons que dans un test statistique classique, après avoir fixé un seuil d'erreur  $\alpha$  : on conclut en comparant la p-valeur calculée à partir des données à  $\alpha$  (p-valeur  $< \alpha \rightarrow$  rejet de  $H_0$ ). Pour les tests incertains,  $H_0$  est rejetée si les deux bornes de la p-valeur sont inférieures à  $\alpha$ , acceptée si elles sont toutes deux supérieures à  $\alpha$ , et si le seuil est à l'intérieur de la p-valeur imprécise, aucune conclusion n'est formulée. Dans [9] Hesamian et Taheri considèrent un degré d'acceptation ou de rejet de  $H_0$ .

Parmi tous les tests cités dans le Tableau 1, les quatre derniers sont les mieux adaptés à la problématique du CHU puisqu'ils sont conçus pour comparer des données appariées et les répétitions de mesures sont effectuées par les mêmes patients au CHU. On pourra néanmoins appliquer des tests prévus pour données non-appariées qui seront alors moins puissants (des écarts plus grands seront nécessaires pour rejeter  $H_0$ ).

En considérant deux ensembles de données imprécises appariées  $[A] = \{a_1, \dots, a_n\}$  et

$[B] = \{b_1, \dots, b_n\}$ , le but de ces tests est de conclure quant à la significativité des différences de distribution entre  $[A]$  et  $[B]$ .

Dans [6], Destercke et Strauss étendent le test de Kolmogorov-Smirnov aux données intervalles pour comparer des distributions de pixels sur une image. Dans notre cas, tout non-rejet de l'hypothèse nulle impliquerait donc l'absence de différence de pression moyenne significative entre les deux types de matelas. L'idée de cette approche est de considérer une statistique de test imprécise sous forme d'intervalle  $[\underline{KS}([A], [B]), \overline{KS}([A], [B])]$  calculé à partir d'ensemble d'intervalles. De manière à réduire la complexité calculatoire, les auteurs proposent d'approximer les bornes de la statistique de test imprécise par les valeurs extrêmes de  $P$ -boîte  $[\underline{KS}([A], [B]), \overline{KS}([A], [B])]$  directement calculées à partir des bornes des éléments de  $[A]$  et  $[B]$  (une  $P$ -boîte est un modèle probabiliste imprécis constitué d'une paire de fonctions de répartition, pour plus de détails les concernant l'auteur est renvoyé à [8]). Ils montrent que  $\overline{KS}([A], [B]) = \overline{KS}([A], [B])$  mais que  $\underline{KS}([A], [B]) \leq \underline{KS}([A], [B])$ .

Le test de Mann-Whitney a été étendu aux données imprécises de différentes façons.

Dencœur et al. proposent dans [5] une extension de ce test aux données floues en se basant sur l'ordre partiel induit par les observations imprécises et calculé à l'aide de l'indice de nécessité de domination stricte entre nombres flous. Les données floues peuvent ainsi être ordonnées et l'extension floue de leur ordonnancement permet le calcul d'une statistique de test incertaine dont les extremums peuvent être obtenus par simulation de Monte-Carlo par chaîne de Markov (du fait de la grandeur de l'espace considéré il est illusoire de le parcourir exhaustivement).

Dans [11], Hesamian et Taheri proposent une extension du test de Mann-Whitney en estimant l'ordre entre observations imprécises à l'aide d'outils issus de la théorie de la crédibilité [20].

Cette approche donne lieu à un degré d'acceptation (ou de rejet) de l'hypothèse nulle  $H_0$ .

Dans un contexte de données imprécises sous forme d'intervalles, Perolat et al. proposent de calculer les bornes de  $p$ -valeur imprécise par un raisonnement ensembliste donnant lieu à un algorithme de faible complexité [2].

A l'aide d'outils possibilistes, Hesamian et Chachi définissent un degré  $W$  d'ordonnement entre deux nombres flous  $A$  et  $B$  par :  $W\{A > B\} = \frac{Pos(A>B)+Nec(A>B)}{2}$ . Après avoir défini une statistique de test imprécise relative à un niveau de dominance  $\gamma \in [0.5, 1)$

par  $\tilde{T}(A, B, \gamma) = \sum_{i=1}^N 1_{\{W(A>B) \geq \gamma\}}$ , une statistique de test générale est définie par  $\tilde{T} = \{\tilde{T}(A, B, \gamma) \mid \gamma \in [0.5, 1)\}$  et des bornes inférieures et supérieures de  $p$ -valeurs sont enfin calculées en considérant les extremums des ensembles  $\{P(\tilde{T}(A, B, 0.5) \geq T) \mid T \in \tilde{T}\}$  et  $\{P(\tilde{T}(A, B, 0.5) \leq T) \mid T \in \tilde{T}\}$ .

Comme explicité par Kahraman et al., toute extension du test de Wilcoxon sera essentiellement dépendante de la manière d'ordonner les observations imprécises [12]. Les auteurs proposent un ensemble d'ordonnement potentiels plus ou moins complexes.

Vaidyanathan étend le test de Wilcoxon aux données floues dans [17] en se basant lui aussi sur la théorie de la crédibilité [14]. En considérant les espérances crédibilistes des différences entre  $[A]$  et  $[B]$ , il se ramène à un contexte précis qui engendre donc une  $p$ -valeur précise.

Taheri et Hesamian proposent une extension du test de Wilcoxon aux données floues par une approche impliquant un ordre entre nombres flous basé sur leur  $\alpha$ -coupes [9]. Pour une comparaison de deux échantillons imprécis, les auteurs proposent un degré de préférence entre hypothèse nulle (égalité) et alternative.

## 5 Proposition de modélisation incertaine à partir de mesures répétées

Soit  $i \in \{1, \dots, n\}$ , on note dorénavant  $x_{1,i}, x_{2,i}, y_{1,i}$  et  $y_{2,i}$  les quatre mesures pour le  $i$ -ième patient, les deux valeurs en  $x$  concernant le premier matelas et en  $y$  pour le deuxième. On suppose qu'elles sont ordonnées, on a  $x_{1,i} < x_{2,i}$  et  $y_{1,i} < y_{2,i}$ .

Cette section présente trois modèles incertains  $\hat{I}, \Pi$  et  $m$  à mettre en œuvre pour convertir les données précises et répétées du CHU respectivement en intervalles, sous-ensembles flous et en fonctions de croyance. Un modèle incertain est ici une application de  $\mathbb{R}^+ \times \mathbb{R}^+$  dans l'espace des intervalles, des sous-ensembles flous ou des fonctions de croyance, e.g.  $(x_{1,n}, x_{2,n}) \mapsto [x_{1,n} - 0.5, x_{2,n} + 0.5]$ .

Les trois types de modèles proposés dans cette section sont des tentatives de projection de l'incertitude issue de la non-répétabilité des données dans des espaces plus large et flexibles que l'ensemble des nombres réels. Ils partagent un même objectif qui est d'étendre les valeurs des mesures répétées vers des supports plus larges que les simples intervalles  $I_i = [x_{1,i}, x_{2,i}]$ . En effet, si on suppose que la non-répétabilité des données est due à une fiabilité partielle des capteurs, les valeurs en dehors de  $I_i$  mais proche ont une réelle pertinence. On définit ainsi le paramètre  $\beta_i$  traduisant l'amplitude cette extension. On suppose que  $\beta_i$  dépend du niveau de fiabilité des données traduit par les écarts  $\Delta_i = |x_{1,i} - x_{2,i}|$  et par l'indice de concordance générale  $\rho$  calculé sur l'ensemble des patients selon l'Equation (1). En partant des hypothèses suivantes :

1. il est d'autant moins probable que la vraie valeur de pression soit dans  $I_i$  que  $I_i$  est étroit, l'extension  $\beta_i$  à effectuer dans ce cas là doit être plus grande que dans le cas où  $I_i$  est large.
2. plus l'indice de concordance  $\rho$  est élevé, plus les données sont fiables et donc

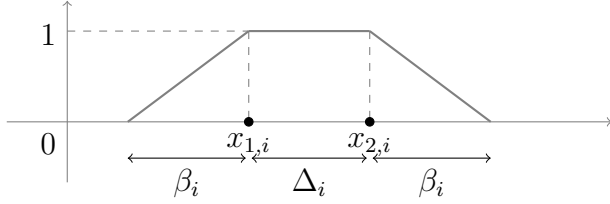


Figure 4 – Modélisation possibiliste

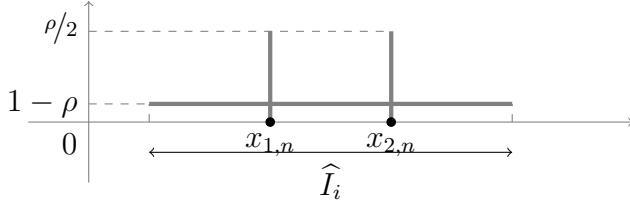


Figure 5 – Modélisation crédibiliste

plus petit doit être  $\beta$ .

on aboutit aux contraintes  $\frac{\partial \beta}{\partial \Delta} \leq 0$  et  $\frac{\partial \rho}{\partial \Delta} \leq 0$ . Une solution à ce problème est :

$$\beta_i = \frac{K}{\Delta_i + \rho + 1} \quad (2)$$

où  $K$  est une constante qu'on fixera empiriquement. A partir de cette proposition on définit les trois modèles explicités dans le Tableau 2.

Tableau 2 – Modèles incertains

$\hat{I}_i$	$= [x_{i,1} - \beta_i, x_{i,2} + \beta_i]$
$\Pi_i$	$= (x_{i,1} - \beta_i, x_{i,1}, x_{i,2}, x_{i,2} + \beta_i)$
$m_i$	$: \begin{cases} m_i(\{x_{i,1}\}) = \rho/2 \\ m_i(\{x_{i,2}\}) = \rho/2 \\ m_i(\hat{I}_i) = 1 - \rho \end{cases}$

Les Figure 4 et 5 représentent les modèles proposés. L'intervalle  $\hat{I}_i$  étend bien l'incertitude issue de la variabilité observée des mesures en dehors de l'intervalle naturel  $I_i$  formé par les 2 répétitions. L'idée de la distribution de possibilité  $\Pi_i$  est de diminuer progressivement le degré de certitude ou de possibilité en dehors de  $I_i$  et de le maintenir constant et maximal dans  $I_i$ . La fonction de croyance  $m_i$  quant à elle, se base sur les informations disponibles en attribuant un même poids aux mesures et un poids à  $\hat{I}_i$  qui sera ici considéré comme le support et servira d'ensemble focal représentant l'ignorance. Ces poids sont fixés de façon globale par le degré

de concordance  $\rho$  des données répétées, l'incertitude aléatoire est ici capturée par le choix de l'ensemble focal  $\hat{I}_i$ .

## 6 Mise en oeuvre des tests sur données incertaines

Après avoir implémenté et testé l'ensemble des approches présentées en Section 4, seules les méthodologies de Destercke-Strauss et de Pérolat et al. ont été retenues d'une part par la simplicité de leur mise en oeuvre et d'autre part car ce sont les seules qui obtiennent des résultats cohérents par simulation de données (à partir de notre implémentation stricte des approches mentionnées en Section 4, en générant aléatoirement l'imprécision). Cette section détaille donc les résultats de ces deux tests incertains sur les données du CHU après modélisation de leur incertitude par les modèles proposés en Section 5. Les codes utilisés pour l'implémentation de l'ensemble des tests incertains décrits dans ce papier sont disponibles sur <https://github.com/lgi2p/uncertainTests>.

Tableau 3 – Résultats de tests incertains appliqués aux données CHU

$p$ -valeur	$0^\circ$	$30^\circ$
Destercke-Strauss	$[0.52, 0.76]$	$[0.50, 0.67]$
Pérolat et al.	$[0, 1]$	$[0, 1]$

Le Tableau 3 présente les résultats obtenus par les tests incertains de Destercke-Strauss (Kolmogorov-Smirnov) et de Pérolat et al. sur les données du CHU après modélisation sous forme d'intervalles  $\hat{I}_i$  comme proposé en Section 5 dans le tableau 2. On conclut donc qu'à  $0^\circ$  comme à  $30^\circ$ , le test de Destercke-Strauss aboutit à une  $p$ -valeur incertaine dont les deux bornes sont au delà de tout niveau de significativité acceptable, ce qui implique une absence de différence significative entre les deux types de matelas. Le tests de Pérolat et al, quant à lui ne permet pas de conclure car il aboutit à des  $p$ -valeurs totalement incertaines (*i.e.* égale à  $[0, 1]$ ) qui ne contiennent donc aucune information.



En l'absence de données supplémentaires, l'approche de modélisation présentée dans ce papier associée aux deux méthodologies de tests incertains implémentées ne permet pas de conclure à une quelconque différence significative entre les 2 types de matelas, le moins cher pourra donc être privilégié par le CHU.

## 7 Conclusion

Cet article présente un travail préliminaire destiné d'une part à proposer une approche de modélisation incertaine dans un contexte de non-reproductibilité des données et d'autre part à uniformiser un certain nombre d'extensions de tests statistiques non-paramétriques aux données incertaines. Des scripts permettant l'implémentation des approches considérées sont rendus disponibles de manière à pouvoir par la suite les compléter et les améliorer.

En termes de perspectives, la réalisation de tests statistiques non-paramétriques à partir de données incertaines crédibilistes reste un problème non-résolu qui pourrait permettre le développement de nouveaux tests incertains prenant en compte un niveau de connaissance supplémentaire relativement à l'incertitude souvent inhérente aux données réelles. Sur le plan hospitalier, les décisions prises par la Haute Autorité de la Santé en termes d'achat de matelas sont pour l'instant restreintes à l'étude d'incidence des escarres. Cet article permet le traitement de données non-reproductibles et contribue ainsi à la prévention des escarres qui surviennent lentement et sont donc difficilement observables.

## Références

- [1] T. Bayes, R. Price, J. Canton. An Essay Towards Solving a Problem in the Doctrine of Chances. *Royal Society of London*, 2015.
- [2] J. Perolat, I. Couso, K. Loquin, O. Strauss. Generalizing the Wilcoxon rank-sum test for interval data. *International Journal of Approximate Reasoning*, 56 : 108-121, 2015.
- [3] A-P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38 : 325-339, 1967.
- [4] T. Denœux. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE transactions on systems, man, and cybernetics*, 25(5) : 804-813, 1995.
- [5] T. Denœux, M-H. Masson, P-A Hébert. Non-parametric rank-based statistics and significance tests for fuzzy data. *Fuzzy Sets and Systems*, 153(1) : 1-28, 2005.
- [6] S. Destercke, O. Strauss. Kolmogorov-Smirnov Test for Interval Data. *Information Processing and Management of Uncertainty (IPMU)*, 2014, pp.416-425.
- [7] D. Dubois, H. Prade. Fuzzy Sets and Systems : Theory and Applications. *New York : Academic Press*, 1980.
- [8] S. Ferson, L. Ginzburg, V. Kreinovich, D. M. Myers, and K. Sentz. Constructing probability boxes and dempster-shafer structures. *Technical report, Sandia National Laboratories*, 2003.
- [9] S. M. Taheri, G. Hesamian. A generalization of the Wilcoxon signed-rank test and its applications. *Statistical Papers*, 54(2) : 457-470, 2012.
- [10] G. Hesamian, J. Chachi. Fuzzy Sign test for imprecise quantities : A p-value approach. *Journal of Intelligent and Fuzzy Systems*, 27(6) : 3159-3167, 2014.
- [11] G. Hesamian, S.M. Taheri. Linear Rank Tests for Two-Sample Fuzzy Data : A P-Value Approach. *Journal of Uncertain Systems*, 7(2) : 129-137, 2013.
- [12] C. Kahraman, C.E. Bozdog, D. Ruan, A. F. Özok. Fuzzy sets approaches to statistical parametric and nonparametric tests. *International Journal of Intelligent Systems*, Special Issue : Intelligent and Soft Computing Techniques for Information Processing, 19(11) : 1069-1087, 2004.
- [13] P-S. de Laplace. Théorie analytique des probabilités. *Ve Coursier*, 1814.
- [14] Uncertainty theory. *Springer, Berlin*, 2007.
- [15] G. Shafer. A mathematical theory of evidence. *Princeton-London : Princeton University Press. XIII*, 1976.
- [16] N. Sutton-Charani, S. Destercke, T. Denœux. Learning decision trees from uncertain data with an evidential EM approach. *International Conference on Machine Learning and Applications (ICMLA 2013)*, 2013.
- [17] V.S. Vaidyanathan. Wilcoxon signed rank test for imprecise observations. *IOSR Journal of Mathematics*, 10(2) : 55-59, 2014.
- [18] R. Viertl. *Statistical Methods for Fuzzy Data*. Copyright © 2011 John Wiley & Sons, Ltd, Wiley Series in Probability and Statistics, 2011.
- [19] P. Walley. Statistical reasoning with imprecise probabilities. *Chapman and Hall*, 1991.
- [20] A-W. Whitney. The theory of experience rating. *Proceedings of the Casualty Actuarial Society*, 4 : 274-292, 1918.
- [21] L-A. Zadeh. Fuzzy sets as a basis for theory of possibility, *Fuzzy Sets and Systems*, 1 : 3-28, 1978.